****

MPSRCH (TM)

****

Release 3.1A John F. Collins, Biocomputing Research Unit.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat May 13 08:46:40 2000; MasPar time 4.48 Seconds
317.408 Million cell updates/sec
Tabular output not generated.

Title: >US-09-331-631-22
Description: (25-84) from US09331631.pep
Perfect Score: 459
Sequence: 1 EDDNHHHHGGHKSGQCVRRC.....EKRQERSRHEADDRSGEGSS 60

Scoring table: PAM 150
Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq35
1:geneseqp

Statistics: Mean 24.771; Variance 107.255; scale 0.231

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	459	100.0	593	1	W62835	Zea mays antimicrobial	8.76e-34
2	180	39.2	637	1	W62837	Hordeum vulgare antimi	1.10e-07
3	121	26.4	625	1	W62830	Macadamia integrifolia	1.48e-02
4	115	25.1	666	1	W62828	Macadamia integrifolia	4.69e-02
5	109	23.7	666	1	W62829	Macadamia integrifolia	1.47e-01
6	102	22.2	432	1	W93954	Human regulatory molec	5.49e-01
7	96	20.9	590	1	W62832	Gossypium hirsutum ant	1.67e+00
8	91	19.8	33	1	W62836	Zea mays antimicrobial	4.18e+00
9	91	19.8	450	1	W46606	Tyrosine kinase associ	4.18e+00
10	89	19.4	186	1	W26536	Trypanosoma cruzi anti	6.00e+00
11	88	19.2	388	1	Y04998	Mycobacterium species	7.19e+00
12	87	19.0	355	1	W05398	Human clone 56 protein	8.61e+00
13	87	19.0	524	1	W32096	Miniature swine retrov	8.61e+00
14	86	18.7	434	1	R96419	Peptide fragment of N-	1.03e+01
15	86	18.7	567	1	W88788	Polypeptide fragment e	1.03e+01
16	86	18.7	797	1	W92641	Murine PGC-1 protein.	1.03e+01
17	85	18.5	314	1	W88499	Human stomach carcinom	1.23e+01
18	85	18.5	525	1	W62831	Theobroma cacao antimi	1.23e+01
19	85	18.5	566	1	R20181	Sequence encoded by 67	1.23e+01
20	82	17.9	660	1	P82921	B subunit of human Fac	2.10e+01
21	82	17.9	898	1	W31853	Mycobacterium tubercul	2.10e+01
22	81	17.6	444	1	W90340	G. max truncated SBP1	2.51e+01
23	81	17.6	524	1	W90339	G. max SBP1 protein.	2.51e+01
24	81	17.6	1148	1	Y07087	Renal cancer associate	2.51e+01
25	81	17.6	2337	1	W37878	Human calcium channel	2.51e+01
26	80	17.4	147	1	Y04801	Mycobacterium species	2.99e+01
27	80	17.4	233	1	Y04815	Mycobacterium species	2.99e+01
28	80	17.4	288	1	W41574	Human cardiac troponin	2.99e+01
29	80	17.4	288	1	W72759	Recombinant human card	2.99e+01
30	80	17.4	409	1	W90342	G. max truncated SBP2	2.99e+01
31	80	17.4	436	1	W03662	Human 70K U1 snRNP pro	2.99e+01
32	80	17.4	489	1	W90341	G. max SBP2 protein.	2.99e+01
33	80	17.4	614	1	R82630	70K autoantigen, part	2.99e+01
34	80	17.4	1311	1	W72971	Precis coenia patched	2.99e+01
35	80	17.4	1311	1	W52197	Precis coenia (butterf	2.99e+01
36	80	17.4	2237	1	R33550	Sequence of the alpha	2.99e+01
37	79	17.2	303	1	R60054	Dirofilaria immitis pa	3.56e+01
38	79	17.2	467	1	W76427	Rat Glucocorticoid ind	3.56e+01
39	79	17.2	605	1	R72608	Human neuronal calcium	3.56e+01
40	79	17.2	612	1	R72609	Human neuronal calcium	3.56e+01
41	79	17.2	660	1	R72613	Human neuronal calcium	3.56e+01
42	79	17.2	1382	1	W31867	Human metastasis-assoc	3.56e+01
43	79	17.2	1829	1	Y07242	Actin-filament binding	3.56e+01
44	78	17.0	162	1	R37991	Sequence of a new cyto	4.24e+01
45	78	17.0	2339	1	W63141	Human calcium channel	4.24e+01

ALIGNMENTS

RESULT 1
ID W62835 standard; Protein; 593 AA.
AC W62835;
DT 27-OCT-1998 (first entry)
DE Zea mays antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Zea mays.
PN WO9827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI; 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 58-60; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 593 AA;

Query Match 100.0%; Score 459; DB 1; Length 593;
Best Local Similarity 100.0%; Pred. No. 8.76e-34;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 25 EDDNHHHHGGHKSGQCVRRCEDRPWHQRPRCLEQCREEEREKRQERSRHEADDRSGEGSS 84
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 25 EDDNHHHHGGHKSGQCVRRCEDRPWHQRPRCLEQCREEEREKRQERSRHEADDRSGEGSS 84

RESULT 2
ID W62837 standard; Protein; 637 AA.
AC W62837;
DT 27-OCT-1998 (first entry)
DE Hordeum vulgare antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Hordeum vulgare.
PN WO9827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI; 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals

```
PS   Claim 1; Page 60-62; 96pp; English.
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
SQ   Sequence   637 AA;

  Query Match           39.2%;  Score 180;  DB 1;  Length 637;
  Best Local Similarity  33.9%;  Pred. No. 1.10e-07;
  Matches    19;  Conservative   22;  Mismatches  13;  Indels   2;  Gaps   2;

Db      30 DDEDDRRGGHSLQQCVQRCRQERPRYSHARCVQECRDDQQQHGRHEQEEEQGRGRG 85
           ||:  ::|||   |||:|| ::||:  ::||:::||:::::  |:|:   ::  |:
Qy      26 DDNHHHHGGHKSGQCVRRC-EDRPWHQRPRCLEQCREEEREK-RQERSRHEADDRS 79


RESULT    3
ID   W62830 standard; Protein; 625 AA.
AC   W62830;
DT   27-OCT-1998  (first entry)
DE   Macadamia integrifolia antimicrobial protein.
KW   antimicrobial protein; infestation; control.
OS   Macadamia integrifolia.
FH   Key             Location/Qualifiers
FT   Peptide         1. .28
FT                   /note= "signal peptide"
FT   Protein         29. .666
FT                   /note= "mature protein"
PN   WO9827805-A1.
PD   02-JUL-1998.
PF   22-DEC-1997; AU0874.
PR   20-DEC-1996; AU-004275.
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI   Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR   WPI; 98-377279/32.
DR   N-PSDB; V42316.
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
PS   Claim 1; Page 43-45; 96pp; English.
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
SQ   Sequence   625 AA;

  Query Match           26.4%;  Score 121;  DB 1;  Length 625;
  Best Local Similarity  31.7%;  Pred. No. 1.48e-02;
  Matches    19;  Conservative   15;  Mismatches  23;  Indels   3;  Gaps   3;

Db      73 EEYNRQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQREDE 132
           |: |:::   :   || :||: |    |  : | ::| |  |:||| :: | | ::|  |
Qy      25 EDDNHHHHGGHKSGQCVRRCEDRPWHQRP-R-CLEQC-REEEREKRQERSRHEADDRSGE 81


RESULT    4
ID   W62828 standard; Protein; 666 AA.
AC   W62828;
DT   27-OCT-1998  (first entry)
DE   Macadamia integrifolia antimicrobial protein.
KW   antimicrobial protein; infestation; control.
OS   Macadamia integrifolia.
FH   Key             Location/Qualifiers
FT   Peptide         1. .28
FT                   /note= "signal peptide"
FT   Protein         29. .666
FT                   /note= "mature protein"
PN   WO9827805-A1.
PD   02-JUL-1998.
PF   22-DEC-1997; AU0874.
PR   20-DEC-1996; AU-004275.
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI   Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR   WPI; 98-377279/32.
DR   N-PSDB; V42310.
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
PS   Claim 1; Page 34-36; 96pp; English.
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
SQ   Sequence   666 AA;

  Query Match           25.1%;  Score 115;  DB 1;  Length 666;
  Best Local Similarity  35.7%;  Pred. No. 4.69e-02;
  Matches    15;  Conservative   14;  Mismatches  11;  Indels   2;  Gaps   2;

Db     182 EEDNKRDPQQREYEDCRRRCEQQEPRQQH-QCQLRCREQQRQ 222
           |:|| :    :   :| ||||:: |::|: :|  :|||::|:
Qy      25 EDDNHHHHGGHKSGQCVRRCEDR-PWHQRPRCLEQCREEERE 65


RESULT    5
ID   W62829 standard; Protein; 666 AA.
AC   W62829;
DT   27-OCT-1998  (first entry)
DE   Macadamia integrifolia antimicrobial protein.
KW   antimicrobial protein; infestation; control.
OS   Macadamia integrifolia.
FH   Key             Location/Qualifiers
FT   Peptide         1. .28
FT                   /note= "signal peptide"
FT   Protein         29. .666
FT                   /note= "mature protein"
PN   WO9827805-A1.
PD   02-JUL-1998.
PF   22-DEC-1997; AU0874.
PR   20-DEC-1996; AU-004275.
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI   Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR   WPI; 98-377279/32.
DR   N-PSDB; V42311.
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
PS   Claim 1; Page 39-41; 96pp; English.
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
SQ   Sequence   666 AA;

  Query Match           23.7%;  Score 109;  DB 1;  Length 666;
  Best Local Similarity  31.1%;  Pred. No. 1.47e-01;
  Matches    19;  Conservative   16;  Mismatches  21;  Indels   5;  Gaps   4;

Db     114 EEYNRQRDPQQQYEQCQERCQRHETEPRHMQT-CQQRCERRYEKEKRKQQKRYEEQQRED 172
           |: |:::   :   ||  ||:  :  |:| :  | ::| |  |:||| :: | | ::|
Qy      25 EDDNHHHHGGHKSGQCVRRCE--DR-PWHQRPRCLEQC-REEEREKRQERSRHEADDRSG 80

Db     173 E 173
           |
Qy      81 E 81


RESULT    6
ID   W93954 standard; Protein; 432 AA.
AC   W93954;
DT   30-JUN-1999  (first entry)
DE   Human regulatory molecule HRM-10 protein.
KW   Human regulatory molecule; HRM-10; cytostatic activity; immune modulator;
KW   transcription factor; enhancer; cell proliferation stimulation; cancer;
KW   treatment; microarray; detection; diagnosis; cell proliferation disease;
KW   leukemia; lymphoma; myeloma; adenocarcinoma; sarcoma; bladder; bone;
KW   brain; lung; liver; ovary; skin; teratocarcinoma; immune response;
KW   allergy; asthma; diabetes; multiple sclerosis; Grave's disease;
KW   myasthenia gravis.
OS   Homo sapiens.
PN   WO9915658-A2.
PD   01-APR-1999.
PF   22-SEP-1998; U19839.
```

```
PR   23-SEP-1997; US-933750.
PA   (INCY-) INCYTE PHARM INC.
PI   Au-Young J, Bandman O, Guegler KJ, Hillman JL, Lal P,
PI   Shah P, Yue H;
DR   WPI; 99-254710/21.
DR   N-PSDB; X24068.
PT   New human regulatory molecules
PS   Claim 1; Page 69-70; 76pp; English.
CC   This invention describes novel human regulatory molecules (HRM) which
CC   have cytostatic activity and act as immune modulators, transcription
CC   factors or enhancers. The HRMs can be used to stimulate cell
CC   proliferation. Antagonists and agonists of the proteins of the invention
CC   can be used to treat cancer. The encoding nucleic acids can be used in
CC   microarrays to detect polynucleotides (and their expression levels) that
CC   encode HRMs in a biological sample. The HRMs and microarrays can be used
CC   diagnose, treat or prevent cell proliferation diseases especially cancer,
CC   e.g. leukemia, lymphoma, myeloma, adenocarcinoma, sarcoma, cancer of e.g.
CC   bladder, bone, brain, lung, liver, ovary, skin, etc, teratocarcinoma, or
CC   to treat or prevent immune responses e.g. allergies, asthma, diabetes,
CC   multiple sclerosis, Grave's disease or myasthenia gravis.
SQ   Sequence   432 AA;

  Query Match           22.2%;  Score 102;  DB 1;  Length 432;
  Best Local Similarity  39.5%;  Pred. No. 5.49e-01;
  Matches    15;  Conservative    9;  Mismatches  11;  Indels   3;  Gaps   3;

Db    306 RRC-SRSRDHKRSRSRERRRTRSRDRRRSRS-HDRSER 341
          |||  |:: | |:|  |: |   |::|: || |:  :|
Qy     42 RRCEDRPW-HQRPRCLEQCREEEREKRQERSRHEADDR 78


RESULT    7
ID   W62832 standard; Protein; 590 AA.
AC   W62832;
DT   27-OCT-1998  (first entry)
DE   Gossypium hirsutum antimicrobial protein.
KW   antimicrobial protein; infestation; control.
OS   Gossypium hirsutum.
PN   WO9827805-A1.
PD   02-JUL-1998.
PF   22-DEC-1997; AU0874.
PR   20-DEC-1996; AU-004275.
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI   Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR   WPI; 98-377279/32.
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
PS   Claim 1; Page 49-51; 96pp; English.
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
SQ   Sequence   590 AA;

  Query Match           20.9%;  Score 96;  DB 1;  Length 590;
  Best Local Similarity  39.5%;  Pred. No. 1.67e+00;
  Matches    17;  Conservative    9;  Mismatches  14;  Indels   3;  Gaps   3;

Db     43 DCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQRHRPEDPQR 85
          :| |||| | |   :: :| | |: :  || :|:| | |  :|
Qy     39 QCVRRCE-D-RPWHQRPRCLEQCREEEREK-RQERSRHEADDR 78


RESULT    8
ID   W62836 standard; Protein; 33 AA.
AC   W62836;
DT   27-OCT-1998  (first entry)
DE   Zea mays antimicrobial protein.
KW   antimicrobial protein; infestation; control.
OS   Zea mays.
PN   WO9827805-A1.
PD   02-JUL-1998.
PF   22-DEC-1997; AU0874.
PR   20-DEC-1996; AU-004275.
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI   Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR   WPI; 98-377279/32.
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
PS   Disclosure; Page 60; 96pp; English.
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
SQ   Sequence   33 AA;

  Query Match           19.8%;  Score 91;  DB 1;  Length 33;
  Best Local Similarity  45.5%;  Pred. No. 4.18e+00;
  Matches    10;  Conservative    5;  Mismatches  6;  Indels   1;  Gaps   1;

Db     10 QCLRRHEGQPW-ETQECMRRCR 30
          ||:|| | :|| :   |: :||
Qy     39 QCVRRCEDRPWHQRPRCLEQCR 60


RESULT    9
ID   W46606 standard; Protein; 450 AA.
AC   W46606;
DT   30-JUN-1998  (first entry)
DE   Tyrosine kinase associated protein 1.
KW   tyrosine kinase associated protein-1; TKA-1; antibody;
KW   detection; binding agent; diagnosis; treatment;
KW   receptor dimerisation.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Protein         7. .112
FT                   /note= "domain repeated in 146-252 and 346-376"
FT   Protein         146. .252
FT                   /note= "repeated from region 7-112"
FT   Protein         346. .376
FT                   /note= "repeated from region 7-112 "
FT   Domain          7. .89
FT                   /label= " GLGF domain "
FT                   /note= "Protein-protein signalling interactions"
FT   Domain          145. .229
FT                   /label= " GLGF domain "
FT                   /note= "Protein-protein signalling interactions"
PN   WO9801551-A1.
PD   15-JAN-1998.
PF   16-OCT-1996; U16510.
PR   13-JUN-1996; US-666067.
PR   13-JUN-1996; US-665037.
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PI   Seedorf L, Ullrich A;
DR   WPI; 98-101049/09.
PT   Tyrosine kinase associated protein-1 - used for the diagnosis and
PT   treatment of TKA-1 related diseases
PS   Claim 2; Fig 1; 61pp; English.
CC   The Tyrosine kinase associated protein-1 (TKA-1) contains a stretch
CC   of 106 amino acids (aa7-112) which is duplicated  within the
CC   sequence (aa 146-252) with 65% identity, and again (aa346-176) with
CC   47% identity. A binding agent of TKA-1, able to bind to a TKA-1
CC   polypeptide, or the full length TKA-1 protein can both be used in a
CC   method for disrupting or promoting receptor dimerisation.  They can
CC   also be used to identify agents capable of interfering with the
CC   interaction between them.  The TKA-1 antibody can be used for the
CC   detection of TKA-1.  The TKA-1 protein, antibody and binding agent can
CC   all be used in the diagnosis and treatment of TKA-1 related diseases and
CC   conditions.
SQ   Sequence   450 AA;

  Query Match           19.8%;  Score 91;  DB 1;  Length 450;
  Best Local Similarity  42.9%;  Pred. No. 4.18e+00;
  Matches    15;  Conservative    6;  Mismatches  12;  Indels   2;  Gaps   2;

Db    397 WQQDRGRETQRCERESETETERERERHRERQRESE 431
          |:| | |  ::| || | |  :|| ||   :| :|
Qy     49 WHQ-RPRCLEQC-REEEREKRQERSRHEADDRSGE 81
```

```
RESULT   10
ID   W26536 standard; Protein; 186 AA.
AC   W26536;
DT   12-JAN-1998  (first entry)
DE   Trypanosoma cruzi antigen.
KW   Antigen; epitope; vaccine; protective immunity; Chagas disease;
KW   diagnosis; therapy; immunoassay.
OS   Trypanosoma cruzi Tulahean strain C2.
PN   WO9718475-A1.
PD   22-MAY-1997.
PF   14-NOV-1996; U18624.
PR   14-NOV-1995; US-557309.
PA   (CORI-) CORIXA CORP.
PI   Houghton RL,  Lodes MJ,  Reed SG,  Skeiky YAW;
DR   WPI; 97-289413/26.
PT   Diagnosing Trypanosoma cruzi infection by detecting antibodies to
PT   novel antigens - which are useful in vaccines to provide protective
PT   immunity against Chagas' disease
PS   Disclosure; Page 70-71; 110pp; English.
CC   This polypeptide sequence comprises a Trypanosoma cruzi full-length
CC   antigen.  T. cruzi antigens (see W26530-41), and epitope-containing
CC   repeat sequences (see W19094-102, W19079-86 and W26542-44) of
CC   native T. cruzi antigens, have been identified by sequencing
CC   genomic DNA and cDNA clones (see T69151-72) isolated from T. cruzi
CC   expression libraries by screening with sera from infected patients.
CC   The antigens can be expressed in transformed or transfected host
CC   cells.  They can be used in a variety of immunoassays for detecting
CC   T. cruzi infection in a blood, serum, plasma, saliva, cerebrospinal
CC   fluid or urine sample.  They are also useful in vaccines and
CC   pharmaceutical compositions for inducing protective immunity against
CC   Chagas disease.
SQ   Sequence   186 AA;

  Query Match           19.4%;  Score 89;  DB 1;  Length 186;
  Best Local Similarity  32.5%;  Pred. No. 6.00e+00;
  Matches    13;  Conservative   14;  Mismatches  11;  Indels   2;  Gaps   2;

Db    128 RREAEERARREVEERARQEAEELARQESEERARQEAEERA 167
          ||  |:|::::   |  ::  |  |:  :|||:|:||::|:
Qy     42 RR-CEDRPWHQ-RPRCLEQCREEEREKRQERSRHEADDRS 79


RESULT   11
ID   Y04998 standard; Protein; 388 AA.
AC   Y04998;
DT   06-JUL-1999  (first entry)
DE   Mycobacterium species protein sequence 50B.
KW   Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW   hybridisation; detection; vaccine; immunisation; infection.
OS   Mycobacterium sp.
PN   WO9909186-A2.
PD   25-FEB-1999.
PF   14-AUG-1998; F01813.
PR   11-SEP-1997; FR-011325.
PR   14-AUG-1997; FR-010404.
PA   (INSP ) INST PASTEUR.
PI   Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y,
PI   Guigueno A;
DR   WPI; 99-181045/15.
DR   N-PSDB; X34249.
PT   Mycobacterial DNA vectors containing reporter constructs - for
PT   identifying coding or promoter sequences involved in
PT   infection-associated protein expression
PS   Claim 32; Fig 50B; 309pp; French.
CC   Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC   from various Mycobacterium species microorganisms.  The encoding
CC   nucleotide sequences can be used as primers and probes for methods
CC   for detecting and identifying mycobacteria, especially belonging to
CC   the M. tuberculosis complex.  The encoded proteins can be used in
CC   vaccines for immunisation against a bacterial or viral infection.
SQ   Sequence   388 AA;

  Query Match           19.2%;  Score 88;  DB 1;  Length 388;
  Best Local Similarity  35.0%;  Pred. No. 7.19e+00;
  Matches    14;  Conservative   10;  Mismatches  14;  Indels   2;  Gaps   2;

Db     37 RRCQHRQRWRRWPRRCRWRWRMALRRRRGRRTRRTRRNRP 76
          |||:|  |:::|| |  : |   | :|  |:|:   :|:
Qy     42 RRCEDRP-WHQRPR-CLEQCREEEREKRQERSRHEADDRS 79


RESULT   12
ID   W05398 standard; Protein; 355 AA.
AC   W05398;
DT   19-FEB-1998  (first entry)
DE   Human clone 56 protein.
KW   Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
KW   cellular signalling element; cellular structural element; malignancy;
KW   protein identification; functional domain; protein screening;
KW   cellular signal transduction process.
OS   Homo sapiens.
PN   WO9631625-A1.
PD   10-OCT-1996.
PF   04-APR-1996; U04454.
PR   03-APR-1996; US-630915.
PR   07-APR-1995; US-417872.
PA   (CYTO-) CYTOGEN CORP.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Fowlkes DM,  Hoffman N,  Kay BK,  Mcconnell SJ,  Sparks AB;
DR   WPI; 96-465045/46.
DR   N-PSDB; T39798.
PT   Identifying polypeptide(s) having specific functional domain (esp.
PT   SH3 domain) - comprises detecting selective binding to recognition
PT   unit, regardless of sequence homology
PS   Claim 54; Fig 57; 174pp; English.
CC   W05386-W05403 represent novel human and mouse Src-homology region 3 (SH3)
CC   domain containing proteins that can be used in the method of the
CC   invention. SH3 domain containing proteins play a role in signalling and
CC   structural elements of cells. The method of the invention is for
CC   identifying polypeptides containing functional domains of interest
CC   (especially SH3 domains). The method comprises contacting a multivalent
CC   recognition unit (RU) complex with a number of peptides and identifying
CC   polypeptides having a selective binding affinity for the RU complex. The
CC   method is based on functional similarities and does not rely on sequence
CC   similarities. Prior methods only gave limited success for identifying
CC   proteins which contain an SH3 domain due to the minimal sequence homology
CC   among known SH3 proteins. It has been found that small peptide RUs in
CC   multivalent form have reduced specificity for a given functional domain
CC   compared to monomer RUs. Multivalent RU complexes are particularly suited
CC   to screening for polypeptides containing functional domains that are
CC   similar to, but not identical in sequence to, the original target
CC   functional domain. The new method enables proteins having a common
CC   function to be identified. Identification of novel SH3 proteins will be
CC   useful for a better understanding of cell growth, malignancy, signal
CC   transduction processes, etc. New candidate drugs can be identified, and
CC   their specificities (e.g. pharmacological activities) can be assessed
CC   using the method of the invention.
SQ   Sequence   355 AA;

  Query Match           19.0%;  Score 87;  DB 1;  Length 355;
  Best Local Similarity  37.8%;  Pred. No. 8.61e+00;
  Matches    17;  Conservative   11;  Mismatches  14;  Indels   3;  Gaps   3;

Db     53 RRLEEKRWAEEAQRQLEQERRE-RELREAARREQRYQEQGGEASP 96
          || |:: |  :  | ||| | | || |:   | :  ::::||:|:
Qy     42 RRCEDRPW-HQRPRCLEQCREEEREKRQERSRHEA-DDRSGEGSS 84


RESULT   13
ID   W32096 standard; Protein; 524 AA.
AC   W32096;
DT   09-FEB-1998  (first entry)
DE   Miniature swine retrovirus GAG protein.
KW   Retrovirus; porcine; GAG protein; xenotransplantation; infectious;
```

KW provirus; organ transplant; donor; activated virus; PCR.
OS Porcine retrovirus.
FH Key Location/Qualifiers
FT Protein 1. .524
FT /label= GAG_protein
PN WO9721836-A1.
PD 19-JUN-1997.
PF 13-DEC-1996; U19680.
PR 14-DEC-1995; US-572645.
PA (GEHO) GEN HOSPITAL CORP.
PI Fishman JA;
DR WPI; 97-332804/30.
DR N-PSDB; T74884.
PT New nucleic acid from porcine retro:viruses - used for detecting
PT viruses in transplant or other tissue and for assessing risk of
PT transmitting infection to graft recipient
PS Claim 22; Fig 3; 128pp; English.
CC This is a porcine retrovirus from miniature swine containing the coding
CC region for a putative viral GAG protein. This sequence and PCR fragments
CC generated from the sequence (see T74812-T74882) could be used to screen
CC organs for porcine retroviruses prior to xenotransplantation.
CC Transplantation can increase the likelihood of retroviral activation if
CC intact and infectious proviruses are present. The porcine retroviral
CC sequence can be used to generate probes to determine the level (e.g.
CC copy number) of intact (i.e. potentially replicating) porcine provirus
CC sequences in a strain of xenograft transplantation donors. It can be
CC used to detect mutations, genetic lesions or viral recombinants and
CC also to determine the histological localisation of activated retrovirus.
CC Using Polymerase Chain Reaction DNA Quantitation (PDQ) on blood
CC mononuclear cells, infectivity titration and susceptibility testing can
CC be performed. Ultimately animal donors without intact porcine retroviral
CC sequences or a lower copy number of viral elements could be selected.
SQ Sequence 524 AA;

Query Match 19.0%; Score 87; DB 1; Length 524;
Best Local Similarity 61.1%; Pred. No. 8.61e+00;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 423 EEEREQRKERERE EREER 440
||||| | || | | ::|
Qy 61 EEEREKRQERSRHEADDR 78

RESULT 14
ID R96419 standard; peptide; 434 AA.
AC R96419;
DT 11-NOV-1996 (first entry)
DE Peptide fragment of N-type calcium channel.
KW Calcium channel; synapse; synaptic vesicle; presynaptic; syntaxin;
KW synaptosome; neuronal cell death; ishaemia; stroke; epilepsy;
KW cognitive deficit; inhibition; screening; detection; treatment.
OS Rattus rattus.
FH Key Location/Qualifiers
FT region 718. .1141
FT /note= "Claimed peptide region."
PN WO9615149-A2.
PD 23-MAY-1996.
PF 09-NOV-1995; U14776.
PR 10-NOV-1994; US-337602.
PA (UNIW) UNIV WASHINGTON.
PI Catterall WA, Sheng Z;
DR WPI; 96-259782/26.
PT Screening for presynaptic calcium channel blockers - identifies
PT compounds which inhibit docking of presynaptic vesicles to calcium
PT channels, rather than compounds which inhibit calcium influx
PS Claim 7; Figure 11A; 53pp; English.
CC A method of screening for compounds that inhibit the interaction
CC between presynaptic calcium channels and presynaptic vesicles
CC comprises contacting calcium channel-like peptide with a candidate
CC compound under conditions sufficient to permit binding between the
CC peptide and the candidate compound, where the peptide is able to
CC bind syntaxin or synaptosome associated protein, and then detecting
CC the presence or absence of binding between the peptide and the
CC candidate compound, thereby determining whether the candidate
CC compound bound to the peptide. The method allows for the screening
CC of compounds which inhibit the docking of presynaptic vesicles to
CC calcium channels and which therefore prevent neurotransmitter release
CC by binding to a selected presynaptic calcium channel-like peptide.
CC Isolated compounds may be used in the prevention of neuronal cell
CC death that accompanies cerebral ischaemia. They may also be used in
CC the treatment of stroke, cognitive deficit related to cardiac
CC surgery and neuronal damage caused during acute epileptic episodes.
CC This sequence corresponds to the LII-III loop (amino acids 710-1143
CC of the rat N-type calcium channel.
SQ Sequence 434 AA;

Query Match 18.7%; Score 86; DB 1; Length 434;
Best Local Similarity 40.0%; Pred. No. 1.03e+01;
Matches 16; Conservative 8; Mismatches 14; Indels 2; Gaps 1;

Db 200 RCERSRRHHRRGSPEE--ATEREPRRHRAHRHAQDSSKEG 237
||| :|:| |: ||| |: |::: |:| | ||
Qy 43 RCEDRPWHQRPRCLEQCREEEREKRQERSRHEADDRSGEG 82

RESULT 15
ID W88788 standard; Protein; 567 AA.
AC W88788;
DT 01-MAR-1999 (first entry)
DE Polypeptide fragment encoded by gene 36.
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN WO9854963-A2.
PD 10-DEC-1998.
PF 04-JUN-1998; U11422.
PR 18-DEC-1997; US-070923.
PR 06-JUN-1997; US-048877.
PR 06-JUN-1997; US-048881.
PR 06-JUN-1997; US-048884.
PR 06-JUN-1997; US-048893.
PR 06-JUN-1997; US-048896.
PR 06-JUN-1997; US-048899.
PR 06-JUN-1997; US-048915.
PR 06-JUN-1997; US-048949.
PR 06-JUN-1997; US-048964.
PR 06-JUN-1997; US-048972.
PR 06-JUN-1997; US-049020.
PR 06-JUN-1997; US-049375.
PR 05-SEP-1997; US-057628.
PR 05-SEP-1997; US-057635.
PR 05-SEP-1997; US-057644.
PR 05-SEP-1997; US-057647.
PR 05-SEP-1997; US-057650.
PR 05-SEP-1997; US-057661.
PR 05-SEP-1997; US-057667.
PR 05-SEP-1997; US-057761.
PR 05-SEP-1997; US-057764.
PR 05-SEP-1997; US-057770.
PR 05-SEP-1997; US-057775.
PR 05-SEP-1997; US-057778.
PR 06-JUN-1997; US-048875.
PR 06-JUN-1997; US-048878.
PR 06-JUN-1997; US-048882.
PR 06-JUN-1997; US-048885.
PR 06-JUN-1997; US-048894.
PR 06-JUN-1997; US-048897.
PR 06-JUN-1997; US-048900.
PR 06-JUN-1997; US-048916.
PR 06-JUN-1997; US-048962.

PR 06-JUN-1997; US-048970.
PR 06-JUN-1997; US-048974.
PR 06-JUN-1997; US-049373.
PR 05-SEP-1997; US-057584.
PR 05-SEP-1997; US-057629.
PR 05-SEP-1997; US-057642.
PR 05-SEP-1997; US-057645.
PR 05-SEP-1997; US-057648.
PR 05-SEP-1997; US-057651.
PR 05-SEP-1997; US-057662.
PR 05-SEP-1997; US-057668.
PR 05-SEP-1997; US-057762.
PR 05-SEP-1997; US-057765.
PR 05-SEP-1997; US-057771.
PR 05-SEP-1997; US-057776.
PR 06-JUN-1997; US-048876.
PR 06-JUN-1997; US-048880.
PR 06-JUN-1997; US-048883.
PR 06-JUN-1997; US-048892.
PR 06-JUN-1997; US-048895.
PR 06-JUN-1997; US-048898.
PR 06-JUN-1997; US-048901.
PR 06-JUN-1997; US-048917.
PR 06-JUN-1997; US-048963.
PR 06-JUN-1997; US-048971.
PR 06-JUN-1997; US-049019.
PR 06-JUN-1997; US-049374.
PR 05-SEP-1997; US-057627.
PR 05-SEP-1997; US-057634.
PR 05-SEP-1997; US-057643.
PR 05-SEP-1997; US-057646.
PR 05-SEP-1997; US-057649.
PR 05-SEP-1997; US-057654.
PR 05-SEP-1997; US-057666.
PR 05-SEP-1997; US-057760.
PR 05-SEP-1997; US-057763.
PR 05-SEP-1997; US-057769.
PR 05-SEP-1997; US-057774.
PR 05-SEP-1997; US-057777.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA,
PI Fan P, Feng P, Ferrie AM, Fischer CL, Florence C,
PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DW,
PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM,
PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
DR WPI; 99-059865/05.
DR N-PSDB; V84446.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Disclosure; Page 30; 772pp; English.
CC The invention relates to nucleic acid sequences (V84411 to V84633)
CC encoding human secreted proteins (W88534 to W88756). The secreted protein
CC gene sequences are deposited with the ATCC under deposit numbers ATCC
CC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC cells comprising recombinant vectors containing the nucleic acid
CC sequences are used for the recombinant production of the secreted
CC proteins. The polynucleotide and amino acid sequences are useful for are
CC useful for preventing, treating or ameliorating medical conditions e.g.
CC by protein or gene therapy. Pathological conditions can be also diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, neurodegenerative
CC disorders, developmental abnormalities and foetal deficiencies, blood
CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung
CC or thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The present sequence represents a polypeptide fragment encoded by a
CC gene of the invention (see descriptor line for gene number).
SQ Sequence 567 AA;

Query Match 18.7%; Score 86; DB 1; Length 567;
Best Local Similarity 36.8%; Pred. No. 1.03e+01;
Matches 14; Conservative 10; Mismatches 12; Indels 2; Gaps 2;

```
Db    497 ERNRQLEREKRREHSRERDRERERERERDRGDRDRDRE 534
          :| :: :| :  |: || :||: :|| |  :| ||  |
Qy     46 DRPWH-QRPRCLEQCREEEREKRQERSRHEAD-DRSGE 81
```

Search completed: Sat May 13 08:46:48 2000
Job time : 8 secs.